ARI Research Note 99-06

Human Performance in Simulation Workshop

Edgar Johnson, Frank Moses, and Joseph Psotka
U.S. Army Research Institute

Research and Advanced Concepts Office
Michael Drillings, Chief

November 1998

U.S. Army Research Institute
for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

A Directorate of the U.S. Total Army Personnel Command

**EDGAR M. JOHNSON
Director**

Technical review by

Michael Drillings

NOTICES

DISTRIBUTION: This Research Note has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This Research Note may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this Research Note are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) November 1998	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) July 30-31, 1997
4. TITLE AND SUBTITLE: **Human Performance in Simulation Workshop.**		5a. CONTRACT OR GRANT NUMBER IN_HOUSE
		5b. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Edgar Johnson, Frank Moses, Joseph Psotka (ARI)		5c. PROJECT NUMBER
		5d. TASK NUMBER
		5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Research Note 99-06
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		
13. SUPPLEMENTARY NOTES		

14. ABSTRACT *(Maximum 200 words)*:

Overview and summaries of presentations at the Human Performance in Simulation Workshop, 30-31 July, 1997.

15. SUBJECT TERMS

Human Performance Performance Live Simulation Virtual Simulation Training Military

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	47	COR: Dr. Michael Drillings 617-8641

Table of Contents

Human Performance in Simulation Workshop
30-31 July, 1997

Overview

1. On July 30-31, 1997, 40 experts, including personnel from the Department of Defense (both military and civilian), industry, and academia participated in a two-day workshop sponsored by the U.S. Army Research Institute and the Institute for Defense Analyses at the request of TRADOC. The workshop assessed the question: Can we represent human performance in simulation better than it currently is done? If so, will it make a difference? The workshop demonstrated that significant knowledge and tools exist to enhance simulations by incorporating human performance factors. However, not all simulations require human performance modeling.

2. The workshop presented state of the art work in live, constructive, and virtual simulation from the perspective of individual, team/group, and command/executive performance. Participants discussed: (a) the importance of collecting and disseminating human performance data from live and virtual simulations, (b) the use of human performance factors in all types of simulations, and (c) the strategies for integrating live, virtual, and constructive simulations into functional systems that serve multiple uses and levels of detail. Human performance data can and should be better organized and modeled to support training, personnel selection, combat operations, advanced doctrine, organizational analysis, systems acquisition and MANPRINT.

3. Capstone Conclusions

a. Existing human performance data could be better used, and made more accessible and influential in the design of simulations. Technology and techniques are available to correct many deficiencies and provide immediate benefits to Force XXI and future Army Warfighting Experiments.

b. Information age military systems typically have a complex human component; current simulations often omit human performance variability. Army as well as joint service simulations and high technology systems need designs that incorporate human performance and address soldier needs. Human performance is widely acknowledged as the most important factor in combat success.

4. Capstone Recommendations

a. Use simulation environments for early system design and development and incorporate the soldier-in-the-loop to reduce the cost of system design and improve training.

b. Integrate human performance factors in simulations that require it, to improve emulation of the unpredictable and unexpected human behavior in military operations.

(1) Make human performance a major component of the design, development, acquisition, and evaluation of simulations as appropriate.

(2) Train decision makers in simulated wargaming environments that exploit the strengths of live, virtual, and constructive simulations to best effect.

c. Develop new models and refine existing models to support direct integration of human performance data; design software to accommodate the data as it becomes available.

(1) Mine existing human performance data and use it in simulations.

(2) Develop and implement embedded data collection techniques in simulation-based environments used for training and rehearsing.

5. Plans for Action

a. Short-term. In a 12-month project, validate the workshop conclusions that systems developed for soldiers and units will be enhanced by incorporating realistic representations of human performance in simulations used to develop the systems. Demonstrate that using simulations for this purpose will provide improved systems quality and usability.

(1) Discussion. Defense Acquisition regulations require the use of models and simulations to reduce time, resources, and risk in the acquisition process (from requirements determination and initial concept exploration to the manufacturing and testing of new systems and related training) and to increase the quality of systems being acquired. However, system developers do not always comply with this regulation and, even when they do, there is no requirement to incorporate human performance in the models and simulations.

(2) Proposed project. Within 12 months after initiation, demonstrate the value of incorporating human performance in simulation by doing so with a system currently being tested, fielded, and already demonstrated as successful. Utilize human-in-the-loop simulation and usability engineering to improve the brigade C4I systems used by the S-2, S-3, and FSO during the recent Task Force AWE at the National Training Center. Interview members of the brigade staff to develop a clear understanding of their interaction with the systems. Additionally, analyze and reconstruct the battles for use in simulation. Execute the project with an integrated process team under the cognizance of the ODUSA(OR) and representatives from TRADOC, DCSPER, ARI, STRICOM, ARL, CECOM, OPTEC, IDA, and the systems contractors. Recommended improvements could be considered for incorporation as upgrades to the current systems.

b. Long-term. ARI will develop a research program to address the key enabling technologies for incorporating human performance components into simulations that require them. The program will be briefed to the AMSEC in February 1998.

6. Workshop presentation topics and details are in the enclosures.

EDGAR M. JOHNSON
Director

Workshop Agenda

Topic	Presenter	Organization
ARI Welcome	Dr. Ed Johnson	ARI
IDA Welcome	GEN (Ret) Larry Welch	IDA
Keynote Address Command and Control	Lt. Gen. Paul Van Riper	MCCDC
Keynote Address Making Sure that Systems are Useful and Usable	Dr. Tom Landauer	U. of Colorado
REQUIREMENTS & ISSUES	Chair: Dr. Ed Johnson	ARI
Architecture High Level Architecture for Modeling and Simulation	Dr. Duke Miller	MIT, LL
Technology TacAir-SOAR: A Synthetic Intelligent Force	Dr. John Laird	U. of Michigan
Combat Human Behavior in Combat Models	GEN(Ret) Dave Maddox	
Commentator	Mr. Walt Hollis	DUSA (OR)
Commentator	Mr. Van Vandiver	CAA
LIVE SIMULATION	Chair: Dr. Robin Keesee	ARL
Individual Live Simulation for the Individual	Mr. Chuck Benton	TSI
Team/Group/Unit The Improvement of Human Representation in Live Simulations for Military Applications	Mr. Jim Madden	IDA
Command/Executive & Staff Human Potential in Live Simulations: The Challenge	BG Pat O'Neal	FORSCOM
Commentator	Dr. Hank Dubin	OPTEC
Commentator	Dr. Rick Morris	TRADOC
VIRTUAL SIMULATION	Chair: Capt(S) Dennis McBride	NMRDC
Individual Team Assessment in Virtual Assessment	Dr. Lewis Johnson	ISI
Individual The Synthetic Forces Program	Cdr. Peggy Feldmann	DARPA
Team/Group/Unit Convergent Simulations: Integrating Deterministic and Interactive Systems	Dr. Bowen Loftin	U. of Houston
Command/Executive and Staff Dynamic Micro-strategies and Cognitive Workload: an Opportunity for Increasing Human Performance via Simulation?	Dr. Wayne Gray	George Mason U
Command/Executive and Staff Representing Command Decision Making in Virtual Simulation	Dr. Lashon Booker	Mitre
Commentator	Dr. Dee Andrews	AFRL

CONSTRUCTIVE SIMULATION	Chair: Mr. Vern Bettencourt	ODUSA(OR)
Individual Rotorcraft Pilot's Associate	Mr. Keith Arthur	ATCOM
Individual Simulations, Intelligence, and Simulated Intelligence	Mr. Darrell Morgeson	LANL
Team/Group/Unit Human Performance in Combat Simulation in Teams	Dr. Andy Belyavin	DERA, UK
Team/Group/Unit Unit Representation in Constructive Models	Dr. George Mastroianni	USAFA
Command/Executive and Staff Affective Computing in Human Performance	Dr. Roz Picard	MIT
Command/Executive and Staff Entertainment Industry Research Directions and Inspirations	Dr. Mike Zyda	NPGS
Commentator	Dr. Al Brandstein	MCCDC
Commentator	Dr. Peter Cherry	VRI

DEVELOP CONCLUSIONS & RECOMMENDATIONS

Co-chairs: Live Simulation Group	Dr. Jack Hiller Mr. Mike Shaler	ODCSPER IDA
Co-chairs: Virtual Simulation Group	Dr. Bowen Loftin Capt(S) Dennis McBride	U. of Houston NMRDC
Co-chairs: Constructive Simulation Group	Dr. Peter Cherry Dr. Al Brandstein	VRI MCCDC

WRAP-UP SESSION	Chair: Dr. Ed Johnson	ARI

Live Simulation Group Presentation
Virtual Simulation Group Presentation
Constructive Group Presentation

Commentator	Dr. Tom Sheridan	MIT
Commentator Getting There from Here - Human Behavior in Existing and Future Simulations of Warfare	Dr. Tom Mastaglio	Old Dominion U.
Commentator	Dr. Paul Berenson	TRADOC

CLOSING REMARKS	Dr. Ed Johnson	ARI

Detailed Workshop Summary

General Conclusions

- Information age military systems typically have a major human component, unlike current simulations where human performance variability often is omitted. Simulations and high technology systems need design and development that incorporate human performance and soldier needs.

- Using human-in-the-loop early in the process of system design and development can avoid costly system performance and training problems.

- Simulations can effectively model early stages of system design and can more easily incorporate the real needs of soldiers and users by cognitive modeling from observation.

- Simulations can assist in improving training, evaluations, decision support systems, and After Action Reviews.

- Simulations can guide development of organizational structures and personnel assignments for using developing equipment.

- Techniques are needed to evaluate the effectiveness of simulations and models to meet specific training and acquisition requirements, objectives, and outcomes.

- Existing human performance data are poorly used, relatively inaccessible, and uninfluential in the design of simulations.

- Human performance data are needed in simulations to represent how the military plan and executes missions.

- New models are not developed and existing models refined to support direct integration of human performance data; software should accommodate data as it becomes available.

- Significant issues and requirements for human performance data in simulations need identification and prioritization; appropriate data sources then can be identified.

- Current techniques are inadequate to measure and collect both objective and subjective human performance data from historical records and from training environments to build required models and simulations.

- An archive, easily accessible to users, is needed that has human performance data sets from training and from historical analyses of combat effectiveness; it could be built using existing Army CATTS and STAARS efforts.

- Human performance in combat is widely acknowledged to be the most important factor in victory and defeat; however, there is only indirect evidence of this in simulations.

- The variability due to human performance factors is left out of many models.

- Beyond the effects of sleep deprivation and some effects of stress, military systems and simulations are not designed to model human performance. We need better methods for simulated representation of human performance variability including psychological, social, and cultural responses to the stresses of war fighting.

- Simulations have begun to incorporate human behavior and cognition, but do not include emotion and personality factors.

Conclusions Related to Particular Issues and Applications

Training

- Data collection techniques embedded in simulation-based environments used for planning, training, and mission rehearsal would vastly improve understanding of human performance.
- Training decision makers to be adept with complex information age systems does not currently take advantage of the strengths of live, virtual and constructive simulations.
- Integrated wargaming in live, virtual, and constructive environments could minimize costs of support personnel.
- Adequate training on new and conceptual equipment in warfighting experiments is essential to develop the right appreciation of its use and effectiveness.
- A better understanding of how learning transfers from simulation experiences to actual operations is essential for improving simulation utility particularly for complex environments and large organizations.
- Understanding the cognitive and emotional analyses of the complex state we now call "will" is important for leadership development. Collapse of the enemy's will constitutes dominance. We need to know the mechanisms and psychological processes that constitute destruction of the enemy's will.
- "One simulation does not fit all" when it comes to representation of human behavior. Effective use of simulations in training can be facilitated by the proper integration of human performance and use of different simulation types (live, virtual, constructive).
- Each type of simulation (live, virtual, constructive) has different strengths and weaknesses to train required tactics, techniques, and procedures. Trainers should match simulation type to training objectives, instructional strategies, and desired learning/training outcomes.
- Enhanced exercise archives with records of plans, preparations, execution, outcomes, observations and After Action Reviews are needed to support accurate analyses and replays in training simulations.

Personnel Selection

- Human performance models and simulations should be developed to provide more valid and effective methods for personnel evaluation given changing military requirements.
- Simulations can and should support a dynamic experiential test-bed for personnel selection that is congruent with future personnel requirements.
- The selection of future military personnel could benefit from methods that are sensitive to changing information system requirements that go beyond current methods which emphasize static knowledge, reasoning, and spatial-relationship tests.
- Current measures of cognitive ability would benefit greatly by supplementing or replacing the use of cognitive abstractions with metrics embedded in concrete experiential situations and realistic simulations.
- Leader selection systems could incorporate assessments of leadership and staff skills, using information age decision supports and simulation-based evaluation metrics.
- Analytic tools and environments are needed to help select and promote excellence in human combat effectiveness based on variables such as cohesion, motivation, courage, and dealing with fear and stress.

Combat Operations

- The quantity of data available to decision makers in today's military operations is complicating the decision-making process; existing decision-making aids need more sophistication to analyze information and to help manage the increasing complexity of combat operations.
- Errors in using command aids result from poor design of human-computer interfaces and lack of system usability considerations that accommodate combat demands. Decision aids should incorporate human information load and management needs in their development and acquisition.
- Digital TOCs do not make effective use of humans-in-the-loop to evaluate and refine prototype tactics, techniques, and procedures (TTPs) pertaining to:

  - communication patterns
  - information utilization
  - battle tracking
  - leadership styles/effectiveness
  - situational awareness

- Decision making exercises using simulation could present realistic future problems and scenarios with unexpected, creative and even brilliant opponents and AARs.

Advanced Doctrine

- Models and simulations that incorporate human performance data could be used effectively for the development and validation of doctrine, and of tactics, techniques, and procedures.

- Current doctrine provides detailed guidance on the performance of procedural tasks, but does not include the effects of human performance variability (both psychological and physiological) and cognitive factors such as skill decay and fatigue.

- Flexible doctrine that properly accounts for variability in human performance is not fully considered in the changing missions of current and future forces.

- The processes of human decision-making are not considered adequately in the development of Command and Control doctrine.

Organizational Analysis

- Testing organizational characteristics or structures in various scenarios is possible through simulation and modeling, although current modeling and simulation does not include sufficient human performance data to support this process effectively.
- Models and simulations with accurate human performance data are not available to assist in the analysis of organizational structures for different military operations.

Systems Acquisition and MANPRINT

- Simulations including user-centered design approaches and rapid prototyping techniques can enhance testing options during the early phases of new system design and development.

- Simulations can provide insight into vehicle and equipment systems' design and acquisition if they allow manipulation of the human-system interfaces and other factors that effect total system usability.

- Simulations of systems that include humans in-the-loop will result in improved designs compared to those that do not by identifying the impact of human performance in both training and operational deployment.

- Adaptability of humans in new environments such as space flight and its simulations help identify how the Army may use multiple technologies to simulate and design new systems.
- Formative evaluations and other user feedback done in simulation environments will facilitate greatly the development of intuitive interfaces that do not interfere with mastery of the system.

SUMMARIES

Lt. Gen. Paul Van Riper

Dr. Tom Landauer

Dr. John Laird

GEN(Ret) Dave Maddox

Mr. Chuck Benton

Mr. Jim Madden

BG Pat O'Neal

Dr. Lewis Johnson

Cdr. Peggy Feldman

Dr. Bowen Loftin

Dr. Wayne Gray

Dr. Lashon Booker

Dr. Dee Andrews

Dr. Duke Miller

Mr. Keith Arthur

Mr. Darrell Morgeson

Dr. Andy Belyavin

Dr. George Mastroianni

Dr. Roz Picard

Dr. Mike Zyda

Dr. Tom Mastaglio

Command and Control

Lieutenant General Paul K. Van Riper, USMC

Our success in the Gulf War in 1991, and the explosive growth of information technologies over the past decade, has resulted in extraordinary claims about the future of war. Claims have been made that "Clausewitz is dead" and that technology will allow us to see everything in the battlespaces of the future--evaporating the "fog" and "friction" of war. In our view these changes do not alter the fundamental nature of war. The microchip has not made Thucydides, Clausewitz, or Mahan irrelevant. In fact, all the trends in modern science, underscore that Clausewitz was right on target regarding the unpredictability of our universe.

The Purpose of Command and Control

Popular literature today is replete with talk about "dominance:" command and control dominance, information dominance, dominant battlespace awareness, dominant battlefield knowledge, dominant maneuver, dominant fires, etc. Command and control is *not* fundamentally about dominance: information is not a medium anyone or any organization can dominate the way the air or sea can theoretically be dominated. Instead, in the Marine Corps view, the fundamental purpose of command and control is: first, to recognize what needs to be done in a situation and second, to see to it that appropriate actions are taken. Command and control is thus essentially about effective decision making and effective execution.

The Command and Control Process: The Observation to Action Loop

We use a simple model to explain the Marine Corps view of the command and control process. Developed by the late Colonel John Boyd, USAF (Retired), it is known as the Observation-Orientation- Decision-Action loop. The observation-action loop essentially describes C2 as a continuous, cyclical process of adaptation to a changing situation.

Engaged in any conflict, we first observe the situation--take in information. Having observed the situation, we next orient to it--make certain assessments, estimates and judgments about the situation and the possibilities. Based on our orientation, we decide what to do. Then we put the decision into action. Having acted we have changed the situation, and so the cycle begins again. The modeling and simulation community has focused on the last step. We can model the movement and interaction of platforms, systems, weapons and units fairly well--at least in a mechanical sense. However, the OODA loop is three quarters cognitive in nature, and very little effort from the M&S community has borne fruit here.

Preparing to Fight in the 21st Century

War contains elements that are both timeless and ever changing. We view the basic nature of war as immutable. It is a violent clash between opposing wills who seek to impose their own will on the other. This interaction occurs cyclically in a series of actions and counteractions between two independent and irreconcilable forces. Our view of the nature of war captures a number of factors including friction, chance, and disorder. Because war is a clash between opposing human wills, the human dimension is central to our views about conflict. Fusing war with intangible factors beyond calculation, prediction, or rational analysis is an absolute necessity, for war is shaped by human nature, the complexities of human behavior, and the limitations of human mental and physical capabilities.

Realizing Command and Control: Creating Better Decision Makers

Without a doubt information is important, but all the information in the world is useless unless it contributes to effective decision making in battle. The U.S. military possesses a plethora of systems today to gather, store and retrieve information. It has numerous programs in place to improve its capacity to manipulate and handle pixels and imagery. Yet information is not knowledge. Our command and control needs to focus, above all else, on providing the combat commander with *understanding* in a form that allows professional judgment and experience to be rapidly applied.

Our philosophy provides a command and control doctrine that accepts war for what it is: an uncertain, tempo-driven, disorderly, and complex phenomenon. It seeks to provide a philosophy of command and control that will allow commanders to make and implement effective military decisions faster than the enemy in any type of conflict, in any setting, on any scale. It relies on *intuitive* decision making based on the latest research in the cognitive sciences to provide the flexibility and responsiveness to deal with uncertainty and generate the tempo that is a key to success in war. It seeks to provide a workable balance among people, procedures, and technology, but recognizes that ultimately there is no substitute for human judgment and understanding.

Our command and control training and educational pursuits reflect the depth of commitment which the Corps' makes towards its most valuable resource, the Marine. Our goal is to equip every Marine with the *thinking* ability to win on the battlefields of the twenty-first century, where the junior enlisted Marine may well need and use more information than a battalion commander does today. The changes are Corps-wide, from the transformation of recruit training to the many steps taken to improve the Corps' entire Professional Military Education program. Other of our initiatives are literally on the "edge of chaos," involving the emerging nonlinear sciences such as chaos and complexity. In addition to staff rides and battle studies, we are also expanding our use of modern computers and simulations to facilitate daily practice in decision making.

Conclusions

The foregoing has provided a macro perspective of how the Marine Corps views "information superiority." Technology permeates every aspect of war, but the science of war cannot account for the dynamic interaction of the physical and moral elements that come into play, by design or by chance, in combat. War will remain predominantly an art, infused with human will, creativity, and judgment. Focusing solely on the technological side will produce "information superiority," but not understanding or the wisdom to know what has to be done. Using "systems" properly within a command and control framework will ensure that our commanders are prepared for the future.

Making Sure that Systems are Useful and Usable

Dr. Thomas K. Landauer

In industrial and business office applications, computer-based systems to help people do their work much better or faster have usually failed to do so. The reason is that systems are usually designed, developed and deployed without the frequent and rigorous testing of their true functionality (helping people do things better, not the number of flops and bytes.) The solution is easy, cheap and effective: test with real humans doing real or simulated jobs before design, during development, and before deployment. Compelling evidence for all three claims will be reviewed. I will argue that the same lack of proper testing and feedback (e.g. not finding out how long it will take a tank commander to type and send an e-mail message BEFORE the system is designed) has plagued many weapons and C3 systems, and that much greater effectiveness can be assured by an engineering discipline called Empirical User-Centered Engineering (UCE). The difficult problem is getting UCE used in an environment where tradition, training, and inclination focuses attention on technical qualities of system and fails to find out what they will do in the hands of soldiers.

TacAir-Soar: A Synthetic Intelligent Force

Dr. John E. Laird

Over the last five years, the Soar/IFOR project of the University of Michigan (UM) and the Information Sciences Institute of the University of Southern California (ISI/USC) has been developing synthetic forces for in-theater air operations. These forces are distinguished by their autonomy. Once they receive the orders, in a form very similar to that received by human pilots, they "fly" their missions according to doctrine without human interruption, except through doctrinally correct interaction over simulated radios. This work has been funded by DARPA. At UM, we have created TacAir-Soar (TAS) for fixed-wing missions. ISI/USC has created RWA-Soar (RAS) for rotary-wing missions. TAS and RAS provide the "minds" of simulated pilots, while ModSAF provides the vehicle dynamics, weapons, sensors and networking infrastructure. In October of 1997, TAS and RAS will participate in the Synthetic Theater of War (STOW-97). For STOW-97, TAS will "fly" all USAF, Navy, Marine, United Kingdom, and OPFOR fixed-wing aircraft missions. RWA-Soar will fly all of the tactical helicopter missions.

TacAir-Soar

TAS entities flies many different missions including defensive counter air, offensive counter air, close-air support, strategic attack, interdiction, airborne early warning, forward air controllers, suppression of enemy air defense, escorts, and tankers. They fly all aspects of these missions, including takeoff, flying in formations, communicating with other entities (using over 200 different message types), in-flight refueling, flying in packages, air-to-air intercepts, air-to-ground attacks, and landing. TAS entities plan and manage the timing and fueling of their missions. They dynamically respond to threats, and changes in missions and weather conditions. Behavior is generated using over 4,800 rules, which are organized hierarchically in terms of doctrine. TAS is implemented in Soar, which has an efficient rule matcher so that it is possible to run multiple independent TacAir-Soar entities on a single machine in real time. Although we have run up to 24 TAS entities on a Pentium Pro, network traffic and interactions with other entities degrade performance enough so that we expect to run between 8 and 12 per machine for STOW-97.

Soar

TacAir-Soar grew out of research on a general artificial intelligence (AI) architecture call Soar. Soar's original purpose was to support the development of AI systems that could use many different problem-solving methods for many different problems. Soar quickly evolved to include the integration of problem solving, planning, learning, and interaction with complex dynamic environments. Much of the inspiration for Soar came from early work on modeling human problem solving by Allen Newell and Herb Simon. Given this background, it is not surprising that Soar was used to develop computational models of human problem solving and learning. All of these models shared the same memory structure, task decomposition, task processing, and learning structure. Based on our initial successes, Allen Newell proposed Soar as a candidate ``Unified Theory of Cognition". Soar has been successfully used to model a wide variety of human behavior; however it is still incomplete in many ways. For example, it lacks a model of the impact of human physiological factors on cognition. However it is one of the few attempts to model a wide variety of psychological effects using a common architecture (John Anderson's ACT-R is another example).

In Soar, all activity is cast as a succession of decisions as to what to do next. The decisions are based on an internal representation of the current situation, which is based on simulated sensors (such as simulated radar, vision, IFF, RWR, and FLIR). To make a decision, a Soar system matches and fires rules to generate preferences for selecting the next "operator". An operator might represent an action as simple as "push the fire button", or as complex as "intercept a bogey". The retrieved preferences are analyzed, and a decision is made for the current best operator. Once the current operator is selected, more rules fire to carry out the actions of the operator. A simple operator will lead to either a new output command being sent (for controlling the plane's controls, weapon systems, sensors, or communication), or some changes to the system's internal state of the system (such as deciding that a bogey is hostile). A complex operator, such as intercepting a bandit, becomes a goal to be achieved through decomposition into simpler operators, which in turn leads to the dynamic construction of a goal hierarchy. Problem solving in goals can lead to the creation of new rules via "chunking." Rules, together with operators and the goal hierarchy provide a smooth integration of reactivity, goal-driven behavior, planning and learning.

Current Research in Advanced Synthetic Forces
In addition to our development for STOW, we are pursuing more advanced research projects with DARPA's Advanced Simulation Technology Thrust (ASTT) and with the Air Force. Under the ASTT program, we are investigating the integration of learning in synthetic forces. In advanced synthetic forces, large knowledge bases are required that are costly and time consuming to generate. In addition, many decisions require a combination of common sense and experiential knowledge that is difficult to extract from an SME using traditional approaches. For example, maintaining situational awareness is one of the primary factors affecting mission success across all branches of the military, but it is extremely difficult to obtain procedures on it from manuals or experts (we've tried). However, trainees learn it, often through a combination of watching others and having an instructor "looking over the shoulder" of a trainee, answering questions, and giving hints during a simulated engagement.

We will use a combination of two techniques: learning by observation and learning by instruction. In learning by observation, a human expert will fly a simulated aircraft and the system will have access to the human's sensory data and actions. The system will attempt to "clone" the behavior of the human by inducing the human's knowledge. We will extend this technique by allowing verbal "annotations". These annotations will be comments, such as, "I'm now turning to intercept the bogey" which will help the system induce the goals that the expert is attempting, allowing the learning to be more general. In learning by instruction, the system will receive instructions from an expert while it is trying to perform the task. The system will use the instructions, not only to perform the task, but also to help it generate internal explanations as to why the instructions are correct. These explanations can lead to generalizations of the instructions instead of just rote learning. Our long-term goal is to build a system that can automatically acquire many of the tactics and techniques that are currently coded by hand.

In addition to our work on learning, we have just started a project with USAF Armstrong Laboratory to model the effects of fatigue on reaction time in a synthetic force. Initially we will explicitly assign reaction times to the various operators and sub-operators used by TacAir-Soar for beyond visual range intercepts. Our hope is that our model will show significant qualitative and quantitative changes in behaviors for different levels of fatigue. More long term, our goal is to develop task-independent computational models of the effects of fatigue on reaction time and error rates, and that these models can be applied to all aspects of our agents behavior.

Recommendations for the development of intelligent forces based on TacAir-Soar work.

1. **Implementation-independent behavior descriptions**. The first step in developing synthetic forces is to precisely define the behaviors that should be encoded. This is extremely challenging and should be done by a team of knowledge engineers and subject matter experts. The description should be in an implementation-independent framework, which does not presuppose how the behaviors will be generated. The danger is that the knowledge engineers will restrict what is included to what is easily represented in their specific scheme. Instead, a reference document should be created that specifies as precisely as possible the desired behaviors; independent of how they will be represented for generation. Needless to say, this document will change during development as both experts and knowledge engineers learn more about what behaviors are required.

2. **Doctrinally correct behavior**. There is a significant temptation to be unsatisfied by synthetic forces, which because of their goal to generate only doctrinally correct behavior, are unrealistic models of human behavior. Research is needed into how to make them show the weaknesses of humans when exposed to extended combat or demoralizing situations, and the strengths of humans in producing behavior that transcends doctrine. However, doctrinally correct behavior is the logical starting point and synthetic forces that could comprehensively generate such behavior would be extremely valuable in many if not all applications. Moreover, because of the breadth and vagueness of doctrine, research is needed in how to go from doctrine specifications to computational representations of behavior.

3. **Interfaces.** The sensory interface should provide the cognitive system with qualitatively the same information that a human has available, such as the objects that are sensed and their basic features. The motor interface should provide actions comparable to those used by the humans, although abstracted so that they directly manipulate objects in the environment. There should be actions to fire a missile, but not to move a finger. It is very tempting to try to simplify behavior generation by providing information unavailable to humans, or by providing complex actions as single motor commands; however both invariably lead to unrealistic behavior. Then again, the problem should not be made any harder than necessary and just as human have automated systems for navigation, identification, and targeting, so should synthetic forces. Thus, TacAir-Soar has simulated versions of many automated systems.

Human Behavior in Combat Models

General David M. Maddox, USA (Ret)

There is no question that people are critical to battle outcome. This has always been true, whether the weapons were clubs, swords, or pistols. However, we developed models and simulations at a time when the questions involved the interaction of large, mechanized forces and the relative effectiveness of major combat systems. So, while these models and simulations under represented and underplayed the actions of soldiers and leaders and ignored the variances in their task performance, these shortcomings may have been acceptable due to the nature of the questions being asked. The result today is that most models and simulations, in particular constructive models, assume perfectly predictable behavior that does not degrade under physical or psychological stress. This predicable perfection is true for individuals, crewmen, commanders and staffs at all echelons.

But the world has changed. We are performing missions with much lower troop densities, we are looking at potential asymmetric actions such as an enemy exploiting the use of dismounted infantry and an urban environment with many non-combatants, and we are in an age where almost any technology is available to a combatant if he has the money with which to purchase it. Models which represent all people the same way, that reflect no differences in training or leadership or response to stress, ignore what may be the most dominant battlefield factors. Simulations which ignored these factors in the seventies and eighties may have been acceptable then. However, adequate representation of human performance and behavior in simulations and models used today in studies of doctrine and tactics, personnel requirements, organizational design, as well as system design alternatives, may provide the difference between future battlefield success and failure.

While the need to adequately represent human performance and behavior in our models and simulations may be understood, the key issue is how to attack this problem. Human performance and behavior — as individuals and as groups and organizations — vary widely and are influenced by many factors, including innate capabilities, training, morale, leadership, cohesion, stress, and fatigue. The ranges of performance and behaviors and the nature of the processes by which levels of performance are established and change are not sufficiently well understood to provide a basis for simulation analogous to what physics and data provide for tank gunnery or sensor performance. Research to establish understanding and provide data is the key to succeeding. It will be a significant task, one which requires careful decisions as to which processes and behaviors should be addressed, in what order or priority and to what levels of detail and confidence. Another issue which must be thought through is how to incorporate the variance of human behavior into our training simulations. One technological enhancement which may assist in this effort is our ability to link constructive, virtual, and live simulation. Our research and early examination of issues involving human behavior may be accomplished through the incorporation of real soldiers and leaders into our simulations.

Live Simulation for the Individual

Mr. Charles Benton

If we take the statement that "All but war is simulation" to heart, then it is obvious that live simulation for the individual is undoubtedly the oldest form of simulation around, dating back thousands of years. Examples of ancient physical simulation abound, including jousting against hay filled dummies and fencing. A long tradition of intellectual simulation for individuals also exists, this makes sense because the most significant tool the individual has to apply to military applications is the mind. Chess certainly could be construed to be a form of simulation that helps to exercise and hone the mental abilities of the individual.

The individual is very unique due to the ability to reconfigure to meet a myriad of different functional requirements. Infantry come in forms ranging from medic to sharpshooter to truck driver to incident commander to computer operator, ad infinitum. Even though live simulation for the individual has been around for millenniums, it tends to be very task specific. A drawback of most present day live simulation capabilities is that they are very special purpose. The simple fact is that we tend not to build simulators for individuals, but instead design simulators to support specific tasks of the individual.

Lastly, this is a workshop about human performance in simulation.... and the term 'performance' implies that a part of the process typically involves some type of measurement or evaluation. For instance, in a live simulation involving pop-up targets, marksmanship is considered essential and is commonly measured using targets or similar capabilities. Thus, part of the overall issue of individual simulators revolves around not only what the individual is doing, but what we measure about that individual's actions.

Many individual tasks, such as navigation through the woods or cleaning a weapon, can be simulated and evaluated quite easily without adding complex technology. Others involve limited technology, such as live-fire training (where blanks are used), but these are still generally straight-forward, mature applications that are well understood.

The individual's potential is leveraged when organized as part of a team, and most individual skills are ultimately aimed at supporting a larger, group goal. Since there is really very little that individuals perform in a 'Lone Ranger' mode, the impact that simulation can have for the individual really boils down to how an individual's environment can be enhanced or augmented to provide a live experience that enables that individual to further his/her skills and knowledge, and how that technology can also be used to assist in evaluating individual performance. What this really means is that the greatest potential for individual live simulation lies in augmenting the live environment to provide a dynamic multifaceted environment.

So what are the implications of these observations? The live environment can be augmented by providing, for example, things for scouts to look for, marksmen to shoot at, and medics to locate. All of these can be provided in low technology forms, for example inflatable OPFOR vehicles for scouts, pop-up targets for marksmen, and surrogate patients for medics. But do these low technology solutions meet the need? In many ways the answer is yes. This type of solution is affordable, and given the flexibility of the individual they represent a cost effective means to enhance the live environment.

But we can also exploit technology to do many things better. We shouldn't do this just because we can, but because it makes a difference in lowering costs or creating a national capability that will provide a real benefit (or better yet, both!). The trend of designing simulators to support specific tasks will continue, even with technological advancements. Indeed, the all purpose immersive environment (a.k.a. the holodeck) is still decades away, and will probably not be realized in our lifetime. But opportunities to enhance the individual's live training environment will become more and more commonplace, and it is important that we understand what is possible and the process to achieve our goals.

Specific technologies are emerging that have direct impact on what is possible. These include new sensor technologies, increased communications capabilities, and truly staggering processing capability in small packages, to name a few. The key to creating effective individual simulators is to find the best mix of subsystem technologies, and to effectively exploit and balance each to achieve an overall capability that is functional, meets user requirements, and is affordable. This leads to unique solutions for each application, and brings us to a common theme which is basic to all good engineering: understanding user requirements.

The process of exploiting technology to meet user requirements is never ending and essentially follows the spiral development process. As an example, a 5-minute video presentation outlining a leading edge capability for individual and teams skills training will be included as part of the presentation. This effort, which is still in the formative stage, is intended to help the workshop focus on how technology can play a role in creating an improved live environment for the individual and assist in measuring individual human performance.

The Improvement of Team, Group, and Unit Human Representation in Live Simulations for Military Applications

Mr. James L. Madden

When asked to address the improvement of human representation in live simulation it's tempting to declare victory as it inherently possesses optimal human representation—the human. There are, however, several issues associated with live simulation that make this a timely subject for discussion.

CURRENT CAPABILITY

The genesis of live simulation in today's Army was GEN Gorman's Board for Dynamic Training that suggested the use of lasers to simulate direct fire weapons systems. This was not an original concept as the British Army was already using the SIMFIRE laser simulator at its Armor school, the test community was using RPMS and lasers (a system similar to that at the NTC), and SAAB had developed a high fidelity duplex laser simulator for evaluating alternative tank main guns. GEN Gorman's genius was his vision of a fully interactive family of laser simulators for all direct fire weapons in a combined arms team whose unit cost would enable them to be utilized within field units rather than just at Service schools. This goal was transformed from concept to system definition by Author D. Little who suggested the use of coded laser transmissions as a means of discriminating between type weapons and the use of Gallium-Arsenide lasers as a means for achieving cost thresholds. Concept definition was turned into reality by Xerox Electro-Optical systems whose simplex (one-way) technical approach won out over ITT's duplex (two-way) technical approach during field evaluations.

At that point the Army was blessed with a great stroke of luck. Kinton, under contract to ARI, came up with the Tactical Engagement Simulation methodology as a basis for improved individual soldier MOS testing. Their premise was that you could not effectively test an individual combatant outside the context of his squad undergoing realistically simulated combat exercises. This methodology called for multiple repetitions of two-sided free play tactical exercises, each followed by an After Action Review. The requisite simulation was provided by a simple telescope affixed to the soldier's rifle, a 3 inch square number affixed to his helmet, and the use of controllers with radios to inform players that they had been hit. The Combat Arms Training Board recognized that this was not merely a new MOS testing methodology, but the essential underpinning for the use of laser simulation devices during tactical exercises. ARI immediately refocused the direction of its research on small units and the Army shortly thereafter fielded the telescope/number simulation system, first as SCOPES for riflemen, and then as REALTRAIN for tanks.

GEN Gorman, the DSCT at TRADOC headquarters by that time, then envisioned the establishment of a fully instrumented National Training Center as an alternative to the largely unsuccessful efforts by many posts throughout the United States to procure additional maneuver areas for emerging longer range weapon systems. This was followed by the introduction of a means to simulate indirect fire, the development of additional instrumented Combat Training Centers, a mobile instrumented range, and most recently, the introduction of a second generation laser simulation device.

With the exception of the development of a means to simulate indirect fires and the introduction of improved technologies, little has really changed with regard to the use of live simulation for military applications over the past 20 years. It is now time for that change to occur.

ISSUES

Two issues suggest where live simulation should go in the future and how it should get there—the evolving battlefield and system acquisition process.

The Evolving Battlefield

The modern battlefield is decentralizing and becoming increasingly technology-based.

When MILES and Tactical Engagement Simulation were introduced, the small unit's combat power was primarily a function of direct fire engagements supplemented with a few remote fire engagements. As one moves into the 21st Century, the small unit's combat power is intended to be a function of remote fire engagements supplemented with a few direct fire engagements. It will generate this combat power not by means of a radio, a rifle, or a TOW, but with a yet to be defined black box of some form, fit, and function that will enable it, in concert with a family of remote sensors, to readily identify targets beyond line of sight, and then immediately engage these targets with remote fire.

The primary focus of live simulation will thus no longer be on direct fire weapons, but on these black boxes. When a small unit goes out to train, how will we generate the sensor input, how will we generate the remote fires, how will we capture the way in which the black box was used in order to focus the AAR's? The easy part is that all of the technologies currently being used for live simulation—lasers, radios, position location systems, terrain databases, and so on— are being incorporated into our next generation weapon systems. The development of live simulations for the 21st Century battlefield should thus be child's play. But for some reason it's not. While there may well be some substantive undertakings going on in this area, the specifications for future operational systems have for the most part been somewhat embarrassing. A typical example was the GEN II Soldier System, an easy target since that program has now been terminated. While it's specifications did indeed call for the use of its inherent technologies, in this case its computer, to support training, the scope of this effort was specified to be the inclusion of the Combat Leader's Guide, color images of foreign weaponry, and the GEN II Technical Manuals into its database. For those of you not familiar with the GEN II Soldier System, it included a radio, computer, laser, and position location capability, all of the major components of a potential live simulation system. It's important to note that this was not the fault of the PM for that system, the fault must lie with the training developers and behaviorists who were advising her.

The Evolving Development Process

Like all great problems the Army has encountered, great solutions are emerging. Foremost, over the long run, is the Simulation, Test, and Evaluation policy working its way through OSD. This transition from a test-fix-test methodology is revolutionary in the way in which simulation and test tools applied in support of the acquisition process are made available for reuse through the life cycle of the system. Models and simulations developed to support the initial phase the STEP acquisition process—Concept Exploration— are refined and utilized to support the three remaining phases of Program Definition, Engineering and Manufacturing Development, and Production/Fielding/ Deployment/Operations. It is specifically intended to develop a joint requirement's, acquisition, and training Test & Evaluation Master Plan at the outset of major programs. With this prospect in hand, one would wonder why the behavioral and training development communities are not breaking down OSD's door to influence this program in their own selfish best interests. There will, of course, be plenty of opportunities in the mid term to implement quick fixes, but the long term, systematic, enhancement of live simulation is clearly best served by getting the Army's premier behaviorists and training developers directly involved with the development of new systems from the outset. The STEP program will enable them to do so.

Human Potential in Live Simulations: The Challenge

BG Pat O'Neal

The views expressed in this paper are those of the author and in no way should be seen as reflecting the official position of the U.S. Army or Department of Defense.

The Problem

Over the past 15 years, the Army's experience in using live simulations to enhance the value of training has proved both the validity of the concept and the need to constantly update our understanding of the factors at work in the live simulation milieu. While we have embraced the concept of live simulation, we have yet to fully explore the factors affecting and effecting the transfer of learning in live simulation. We have yet to take any significant and academically rigorous steps to find the correct combination of factors that, when present in proper proportion, create optimum conditions for the transfer of learning. As a result, the National Training Center (NTC) and most of our home station training is most likely operating at a fraction of potential efficiency. The linkage of this low level of efficiency from the NTC to home station training results from a natural desire on the part of leaders to model the functionality found at the NTC. Therefore, our home station training inherits the natural inefficiency from the NTC.

A Way Ahead

We need to launch an effort to explore technology that can help to document the transfer of learning. In particular, we need to engage academia to study the transfer of complex tasks within large organizations. While we have a great deal of evidence to support the transfer of training in technical skills such as pilot training in virtual and live simulation, we have little evidence for the transfer of learning of complex skills in large organizations involved in land warfare.

We continue to leverage technology to unload the soldier, airman, sailor, and marine. We look for new technological ways to protect the service member, enhance the lethal potential of the service member, enable the soldier to better navigate the battlefield, or improve the commander's ability to command and control forces. Yet, although we know technology can multiply the soldier's potential by a factor of three, we continue to ignore its application for quantifying the effect of training and the associated transfer of learning, both of which could magnify the soldier's potential by a factor of five. We in DoD need to launch a long-term and academically sound study of the transfer of training in live simulation. As we become more "joint" in the way we approach operations, we need to understand the most efficient and effective methods to train large joint forces in complex operations and in live simulation. We also need this empirical data base to properly justify funding for increasingly expensive training.

The intent of developing an ability to create measures of effectiveness, on the transfer of learning is not to grade the commander's ability to perform in the live simulation environment. If we fall victim to the temptation to compare unit performance in the live simulation environment, we will destroy the fabric of leader trust developed over the past 20 years. However, in creating the ability to give the commander some valid and empirical evidence to accompany his intuitive evidence of the transfer of learning (not merely how many times the unit wins or loses), we help every member of the organization to understand entry and exit levels of learning.

We tend to shy away from the study of learning, preferring instead to concentrate on the more tangible and technologically sophisticated aspects of hardware. We now need to strike a balance and begin to concentrate on the rich vein of human potential. It is time to engage the rich fields of Organizational Psychology and Education to better understand how to tap more of our force's human potential in the training environment.

Team Assessment in Virtual Simulations*

Dr. W. Lewis Johnson

The key to effective use of virtual simulation technology is the ability to perform accurate assessments of force behavior. In simulation-based training, for example, one must carefully evaluate the performance of the trainees against the training objectives of the exercise, to determine how well they performed the skills that they were expected to perform. Accurate assessment is critical to the use of virtual simulations in developing new tactics and force configurations for the Army of the future.

This presentation will describe an approach to modeling and analyzing team behavior that has led to the development of new software tools for assessing team performance in virtual simulations. The modeling approach involves recognizing tactically significant situations as they arise during the simulation. Once the situation is interpreted, knowledge of relevant tactics can be brought to bear to recognize the team's actions and collect the necessary data to assess performance, both in real time and in after-action review. This research has resulted in a tool called PROBES that supports observer controllers in assessing simulation exercises.

Motivation

This work is based on observations of how instructors currently manage armor training exercises at Fort Knox, as well as team training in the Navy and the Air Force. Army instructors do an excellent job of managing training exercises, but the tools that they have at their disposal are severely lacking. Instructors are often forced to write notes to themselves on paper as the exercise proceeds. Observer controllers are often observed asking questions about past behavior, like "How long has that unit been in its current position?" It is not always possible for the exercise controller to review the past events in the exercise while at the same time managing the current exercise activities. ModSAF-controlled OPFORs can provide information about their activities that is useful for assessing trainee activities, but only if the exercise controller requests it. Furthermore, this information is not recorded, and is therefore not available during action review.

When one moves to scenarios involving larger numbers of units, each of which is operating semi-independently, the problem of performance monitoring becomes more severe. For example, when the Navy trains engine operations in simulated scenarios it employs entire teams of trainers, one trainer for each trainee. The trainers follow the trainees around, asking them questions as they go. This model is clearly very costly, and should probably be avoided if possible. However, without automated support, it may become increasingly difficult for individual observer controllers to manage and analyze exercises, particularly in the decentralized force configurations envisaged for the future.

PROBES

PROBES provides intelligent assistance to instructors in the form of *intelligent probes* that:

- detect and report events/trends/interpretations relevant to training objectives relevant to training objectives,
- aggregate analysis of multiple entities, to help with the management of complex, multi-player scenarios,
- present the activity to the instructor in both 2D and enhanced 3D presentations, and
- create a record of significant events in the exercise that can be used both in real-time performance assessment and in after action review.

PROBES provides instructors with a coordinated set of presentations. One presentation depicts the current situation of the force in the context of a *situation space.* The situation space enumerates the different types of situations that the soldiers may find themselves in, each of which is associated with different tactical and instructional objectives. PROBES automatically determines which situation is currently in force, by monitoring the activities in the simulation. It uses synthesized speech to notify the instructor of state changes, so that the instructor knows the current situation regardless of which display he is currently looking at.

The exercise analysis log records events and analyses that are likely to be relevant to the current situation. When a trainee platoon initiates Situation Travel (Wedge) [i.e., starts traveling in a wedge formation], PROBES reports when it notices that vehicles in the platoon are in a poor wedge formation. This analysis is performed only when PROBES recognizes that the platoon should be in a wedge. In contrast, when Situation Act on Contact occurs

* This work is supported in part by the DARPA Computer Aided Education and Training Initiative, partly by the Office of Naval Research, and partly by USAF Armstrong Laboratory.

[i.e., action on contact is initiated], PROBES starts reporting that the platoon is still in a wedge when it should be on line. At this point it stops reporting on whether the wedge is in proper alignment, since the wedge formation is no longer appropriate for the current situation.

PROBES monitors not just the movement of the vehicles, via the DIS packet stream, it also accesses state information internal to the vehicle simulations. This allows PROBES to report on the OPFOR's tactics and situation assessments. In the case of the blue forces, PROBES infers their intent based on the current situation and the objectives of the exercise.

As the situation changes, PROBES automatically displays statistics about the unit's performance as appropriate to the current situation. For example, when the platoon is travelling in a wedge formation a diagram appears that shows the following distances and depths of the wedge.

If the instructor wishes to see more information about an observed event, he can obtain it simply by clicking on the event notification in the event log. This causes PROBES to bring up an additional window showing local terrain conditions, damage to the vehicles, etc. This information can help the instructor to evaluate whether or not the trainee behavior is appropriate in the current circumstances. This same capability may be used during after action review.

Finally, we are coupling the PROBES analysis with the stealth display. The instructor can click on an event in the exercise analysis log and see a 3D view of the situation at that time. We are experimenting with ways of annotating the stealth display with relevant analysis data. For example, red and green lines indicate whether or not the vehicles have recently been in each other's line of sight – green means that the following tank could recently see the leading tank, red means that the following tank could not. Sliders indicate the tank's following distance relative to doctrinal norms. A blue ring represents the following tank's current position, and red rings represent the upper and lower limits on following distance. This and other annotations allow the instructor to assess the formation accurately from any view direction.

PROBES is able to analyze exercises because it has access to more information about the engagement than what is available through the DIS packet stream We created an separate event stream to carry the information needed for exercise analysis. In a PROBES-enabled simulation, trainees fight against semi-automated forces that have been specially instrumented to report their tactical decisions. The Puppet Master, the heart of the PROBES system, orchestrates the analysis by sending data requests to the instrumented SAFs, interpreting the data, and displaying the results.

Tracking Team Interactions

In order to track more complex team interactions, it is necessary to model how activities of each individual team member and how they interact with the activities of the other team members. For this purpose the team training capabilities build into the our pedagogical agent architecture Steve (Soar Training Expert for Virtual Environments) may be appropriate. Steve is an autonomous agent that can either play the role of a team member or can advise individual trainees are they participate in team tasks. Steve's knowledge of the task is represented as a network of partial-order plans, describing the activities of the individual team members and the ways in which they depend upon each other. Using this model, Steve can either play the role of a team member or track what other team members are doing. This tracking capability should make it possible to extend PROBES to assess interactive team behavior.

The Synthetic Forces Program

Cdr. Peggy Feldmann, USN

The Synthetic Forces Program is an integral part of the Synthetic Theater of War 97 Advanced Concept Technology Demonstration (STOW-97 ACTD). It has produced, over the last three years, the forces needed to train in a platform-based, seamless, joint synthetic battlespace. Building off Modular Semi-Automated Forces (ModSAF), Synthetic Forces now exists for Army, Navy, Marine Corps, and Air Force platforms, as well as, Opposing Forces and United Kingdom Forces. The construct of these forces include finite state semi-automated forces, rule based multi-echelon command forces, and artificial intelligence pilots. Building decision-makers who plan, replan, and execute higher orders has been a primary goal of the program. This paper provides a description of the Synthetic Forces (SF) used for STOW-97 ACTD with an emphasis on the human aspects. The STOW-97 ACTD will be an integral part of United Endeavor 98-1 (UE 98-1), a United States Atlantic Command (USACOM) sponsored exercise held in late October and early November 1997. Included is a broad overview of how forces are simulated and describes how they move, shoot, communicate, think, and interact.

For the Synthetic Forces development, the initial requirements for the STOW-97 ACTD were extremely ambitious. DARPA had to develop entity level Synthetic Forces for all of the Services suitable for Joint Task Force (JTF) Tier III (JTF Staff) training. To accomplish this within the time and funding available, DARPA chose to build upon the basic architecture of Modular Semi-Automated Forces (ModSAF) to develop separate Synthetic Forces for the Army, Navy, Marine Corps, Air Force and Opposing Force (OPFOR). This proved extremely challenging since we had to modify the underlying architecture while simultaneously developing new forces and capabilities.

In general, the behavior of an individual SAF entity is produced by an integration of individual behaviors, which react to the entity's immediate environment, and unit level behaviors, which coordinate the activities of multiple entities, such as platoons or companies. To increase the fidelity of the representation of the command decision and communication process, STOW developed Command Forces (CFOR). Rather than imposing a specific methodology for representing behaviors, CFOR uses a variety of sophisticated Artificial Intelligence techniques. Underlying these techniques is a common infrastructure which includes utilities for accessing complex terrain information, as well as a common language, CCSIL (Command and Control Simulation Interface Language), to support communication between SAF entities, Command Entities, and C4I systems.

Convergent Simulations: Integrating Deterministic and Interactive Systems

J.A. Bartasis, B.E. Johnston, and R.B. Loftin

Simulation-based training can trace its history through armies the world over. The use of live simulations to prepare war fighters has developed along an historical continuum spanning thousands of years—from the *Armatura* of the Romans, to the *Butokuden* of the samurai class, to the *Quintain* of the Middle Ages, and finally to the National Training Center (NTC) of today. Live simulations have provided the soldier with an opportunity to learn his trade in a "hands-on" arena using a time-tested methodology.

Constructive simulations, in the form of war gaming, may be almost as old as live simulations. The board games of ancient Egypt, China, and the West focused on strategy, pattern recognition, and prediction. The specialization of war gaming came with the German *Kriegsspiel* which established the modern metaphor for the type of simulation that is embodied in today's JANUS system.

Virtual simulations, however, are essentially products of the twentieth century, initially defined by Edwin Link's famous "blue box" that trained thousands of aviators during World War II. Following the industrialization of modern warfare, such virtual simulations have steadily matured to yield the Conduct-of-Fire Trainers (COFTs), both rotary and fixed-wing aircraft simulators, and ship bridge simulators that are now in routine use. While virtual simulations are a modern product they can trace their history to their live simulation predecessors.

The last decade has seen the evolution and growth in scope of Distributed Interactive Simulation (DIS), from the early SimNet to the Close Combat Tactical Trainer (CCTT). The coming of age of DIS has been marked by the convergence of live and virtual simulation, evidenced by trainees' decreased ability to differentiate their participation in training as part of a real or virtual experience. This convergence of simulation does not, at present, extend to constructive simulation. The ability of constructive simulations to interact successfully with virtual and live simulations has been shown to be, at best, incomplete. For example, a 1991 study[1] of JANUS in modeling an NTC MILES field exercise showed "statistically significant attrition differences . . . between a similar NTC MILES and JANUS(A) scenario. . . ."

Diversity of training goals and representation of human interactions among and within live, constructive, and virtual simulations encourages a separation and specialization of resources through the system architecture. Procedural training outcomes are often produced most efficiently by algorithmic computations in constructive and virtual simulations. Human behavior in interactive contexts and the presentation of multiple levels of activities and consequences are not well served by this method.

As simulation-based training is expanded to a larger segment of the war fighting force, this shortcoming will become more pronounced, further separating the effective application of constructive simulations from live and virtual simulations. Such a divergence creates a waste of resources and training effort. It is the purpose of this paper to explore this divergence among resources in an effort to develop methods that will facilitate the convergence of all three simulation types.

Within constructive simulations human interactions are considered as isolated, discrete events that are initiated by goal-seeking behaviors or as challenges implemented to evaluate individual performance. These interactions are captured, monitored by human or computer agents, and provide data that informs the evaluation of human performance. Accurately predicting and modeling human behavior in algorithm-based systems is limited to collective levels of activity. However, this level of modeling can present representation problems when attempts to include the effect of individual actions are made. Other attempts to model human behavior succeed primarily in well-constrained circumstances, as in the recent triumph of IBM's Deep Blue over the human world chess champion. Few would argue, however, that war fighting, in its broadest context, is as rule-driven as chess.

Human behaviors are inherently complex, generally non-linear, and difficult to predict with a high degree of confidence. Multi-disciplinary exploration of the relationships between human behaviors and computer-generated models is underway in an effort to represent the eloquence and complexity of human interactions.

---

[1] D.A. Dryer, "Comparison of the Janus(A) Combat Model to National Training Center (NTC) Battle Data," U.S. Army Training and Doctrine Analysis Command, Monterey, CA, June, 1991 (TRAC-RDM-TR-191).

One promising solution is evident in Peter Wegner's argument[2] for the superiority of interactive systems over algorithms as problem-solving engines. Wegner maintains that algorithms yield outputs that are always determined by their inputs, while interactive systems are history-dependent, capable of learning and adapting through experience. Human behavior is analogous to such interactive systems. Interactive systems, in conjunction with the Army's rich data base of human behavior, have the potential to effectively model the complexities of individual or small group behaviors. Interactive systems also offer the possibility of representing multiple perspectives within the training scenario.

One way that representation of human behavior by interactive systems can be realized for team training is through the use of agent-enriched virtual environments. This technology is characterized by continuous tracking of the human user's behavior in the virtual world. Concurrently, the evolution of artificial intelligence from a reliance on logic-based models to agent-based models provides the military training community with a means of extracting meaning from, and injecting extraordinary training value into, virtual environments. Agents, embedded in virtual environments, can capture human behavior and, relying on an experience-based model derived from human performance data, adapt and learn as the environment is used for training. More importantly, training simulations of this kind optimize the power of each type of simulation, facilitating a convergence of resources to provide the most successful training outcomes.

The vision of future training can be one of the convergence of algorithmic-based simulations for strategic training with live and virtual simulations for improving the performance of individuals and small teams. Improvements in technology and sensory fidelity will guarantee the continued convergence of live and virtual simulation, reducing the need for maintaining large-scale live training sites as these sites become increasingly more difficult and expensive to maintain. These convergent simulations will also be more amenable to adaptation for the growing diversity of war fighting scenarios, including joint operations and special weapons and tactics. The basic elements needed to achieve this convergence of simulations are in hand—complex, real-time virtual environments, layered multi-perspective displays, agent-based architectures, and a growing body of data on human performance. Through coordination of effort and careful investment, convergent simulations can be brought to bear on the training of the Army After Next.

---

[2] P. Wegner, "Why Interaction Is More Powerful than Algorithms," *Communications of the Association for Computing Machinery*, Vol. 40, No. 5, pp. 80-91 (May, 1997).

Dynamic micro-strategies and cognitive workload: An opportunity for increasing human performance via simulations?

Dr. Wayne D. Gray

Preview

Can ten hours spent playing a video game lead to better performance and a higher success rate for Air Force pilot candidates? Based upon a research study (Gopher, Weil & Bareket, 1994) the Israeli Air Force thinks so and requires its pilot candidates to play ten hours of the game Space Fortress. The claim is that during this ten hours, pilot candidates are learning better strategic control of visual attention. Is this a wild claim? Maybe yes, maybe no.

Overview

Recently my research has led me to postulate dynamic micro-strategies that impact cognitive workload or, colloquially, "our ability to do and remember 20 different things at once". The elements of these dynamic micro-strategies take from 30 to 300 msec to occur. The micro-strategies themselves require 300 msec to 3 sec to execute.

These dynamic micro-strategies occur far below the timescale required to play a game of Space Fortress and much below the time it takes to fly a plane or to do anything else of interest in the real-world. Indeed, they are even below the un-real-world tasks studied by human factors psychologists when they study cognitive workload.

In an information rich environment, a dynamic micro-strategy entails a trade-off between continued processing of what you are paying attention to now (the current candidate item) versus finding, attending, and processing the next candidate. I refer to these as a tradeoff between grazing (information processing at one location) versus browsing (moving on to another location).

Issues to consider in dynamic micro-strategies include the following;

- browsing costs
  - cost of locating next candidate
  - cost of reallocating attention to next candidate
  - cost of re-finding current candidate (after moving on to the next candidate)
- grazing costs
  - "sunk" costs of processing current candidate
  - cost of additional processing of current candidate
  - cost of reprocessing current candidate to current level of analysis (after moving from and then re-finding the current candidate)

Mundane example

Your desk computer is down. To complete slides for a presentation that is due tomorrow you have borrowed a friend's Macintosh. Rather than Powerpoint or Harvard Graphics, she uses something called ClarisImpact to prepare her slides. There is no instruction manual; however, you are confident of success as "all these things" do essentially the same thing, and it is simply a matter of finding out how to do what you want with the current system.

To do this task you rely upon three levels of menus. Across the menuBar at the top of the screen is the apple menu plus 8 items specific to ClarisImpact. When the mouse is moved to a menuBar item and the mouse button is held down, a menu drops down. This menu has menuItems. When the mouse is moved to a menuItem and the mouse button is released then the menu disappears and either a palette pops up or the system does something to your presentation. If you select the wrong palette or inadvertently put the system into the wrong state it may take some time (and exploration) to undo this and return to what you were trying to do.

As I start to observe you, you are looking for a function that has to do with formatting text. Initially, your eye movements show that you are looking and very briefly considering all nine of the menuBar items. When you start moving the mouse, you ignore four of the nine and spend most of your time moving among the FILE, EDIT, LAYOUT, and TEXT menuBar items. At each item, you briefly hold the mouse button down, then you release the mouse button and move on to the next menuBar item. You go through this procedure many times. Each time that you pull one of these menus down you study its menu items (graze). However, rather than selecting a menu item right away, you go to (browse) the next pull down menu and study (graze) its items. After several seemingly random iterations you go to the LAYOUT menu, drag your mouse to "RULERS . . ." and mouse up; a rulers palette appears and you continue on.

This common example demonstrates a search among three different sets of menus with different but complementary micro-strategies used at each. The top level entailed a search among the 9 menuBar items. The cost of locating the next candidate was trivial as was the cost of reallocating attention. Both could be accomplished with the same eye movement. On each browse a little grazing was performed but not much, as you quickly moved onto the next menuBar item. Your processing at this level can be characterized as a low-cost progressive deepening. The main result of this processing was to eliminate five of the nine items from further consideration.

The next level involved moving the cursor to a menuBar item and holding the mouse down while your eyes browsed and grazed the menuItems. Here the cost of locating the next candidate includes the cost of moving the mouse between menuBar items (with mouse movements and mouse downs). Although these costs are minor, they are substantially higher than the cost of an eye movement. Finally, reallocating attention to a drop down menu item required an eye movement. The micro-strategies involved at this level include browsing as well as grazing these menuItems.

The third level involves selecting a menuItem from a drop down menu. This selection involves dragging the mouse to the menuItem and releasing the mouse. However, included in the cost of locating the next candidate are the costs of making an error. Unlike the first two levels, browsing the wrong selection has a real cost associated with it. Avoiding this cost is what drives much of the level 2 browsing and grazing behavior. The assertion is that level 2 browsing and grazing is as extensive as it is primarily due to the cost associated with a level 3 false alarm; i.e., deeper level 2 processing helps ensures that the level 3 selection is the best of those available.

Implications for Human Performance in Simulation

My example simplifies but distorts many things. I hope it provides you with an appreciation of the level at which the micro-strategies operate as well as with a sense of their ubiquitousness. It should also serve to provide you with one sense of the dynamic aspect of dynamic micro-strategies; the micro-strategy used on level 2 menu items is partially determined by the costs associated with the level 3 menu selection. On the other hand, the example involved a task where the pacing is entirely under the control of the user; the task itself was not very dynamic. Dynamic micro-strategies are a real-time response to event-driven (or data-driven) tasks such as flying a plane or attempting to classify targets in an air defense scenario. They are the building blocks of more elaborate strategies, those that we consciously choose among. None of these micro-strategies are wrong, per se; however, some are more inefficient in the context of a given task than others.

Micro-strategies are not chosen deliberately, but are selected in real-time, the selection of one micro-strategy vs another is a result of subsymbolic selection processes that are below our conscious awareness (Anderson, 1993). The 300 msec to 3 sec range for dynamic micro-strategies does not leave much time for conscious thought or selection among strategies. It implies that before getting into a situation the user needs to have already acquired the correct micro-strategies and have already acquired selection rules for when to use one strategy vs another.

Our current hypothesis is that choosing inefficient micro-strategies is a common source of cognitive workload and much anecdotal evidence suggests that people are bad at choosing good micro-strategies. Indeed, as this paper was being written I learned of a recent experiment in which different groups of subjects accomplished the same task using either a direct manipulation interface, a natural language interface, or a mixed interface. Of the three groups, the mixed group (i.e., the only group that could choose which modality to use for each part of each interaction) was the worse (J. Gregory Trafton, personal communication, June 23, 1997). Clearly, choosing the correct micro-strategies for each part of each interaction is problematic, not obvious.

If choosing a micro-strategy is not a conscious choice, how can we train people to select the optimal micro-strategies? While it is early days for our research program, we believe the answer will entail some form of simulation. Given that so many micro-strategies exist and given that many different ones will accomplish the same task, learning when to use what dynamic micro-strategy requires getting people to pay attention to features of their environments and outcomes to which they do not normally attend. Perhaps this is what Space Fortress did for the Israel Air Force pilot candidates. Discovery learning is not the answer. The answer requires the creation of a simulation environment in which controlled practice and focused feedback can be provided.

Acknowledgments

The work at George Mason University was supported by a grant from AFOSR (F49620-97-1-0353) to Wayne D. Gray.

References

Anderson, J. R. (1993). Rules of the mind. Hillsdale, NJ: Erlbaum.

Gopher, D., Weil, M., & Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. Human Factors, 36(3), 387-405.

Representing Command Decision Making in Virtual Simulation

Dr. Lashon B. Booker

The DARPA command forces (CFOR) project, part of the Synthetic Theater of War (STOW) program, has implemented the first explicit models of command and control in virtual simulation. The CFOR approach is based on several key ideas: (1) a representation of command and control in terms of the behaviors and information exchanges generated by decision-making entities (2) a common set of structured messages (the Command and Control Simulation Interface Language (CCSIL)) that explicitly represent the C2 information exchanged among entities (both simulated and human participants); (3) an architecture in which representations of command nodes (called command entities) interact with the virtual battlefield through a set of common services; (4) a development strategy that integrates the efforts of multiple developers using a variety of technical approaches to model command decision-making. The CFOR approach has facilitated the development of software command entities that represent the state-of-the-art in modeling command decision making for virtual simulation.

Given the availability of fairly robust models of entity behavior for platforms and small units (e.g. platoon level and below), the CFOR program focused on modeling command decision making at the next higher echelon (e.g., company) where human operators typically intervene to control behavior. Current implementations of command entities have proven to be capable of making autonomous, doctrinally correct command decisions in a variety of tactical situations. They are able to interpret an incoming operations order, plan a course of action to achieve their assigned mission, issue appropriate orders to their subordinates, and control the execution of the subsequent tactical operation. CFOR command entities will play a key role in the STOW ACTD. For example, the combined arms operations of all Army tank and mechanized infantry companies and company teams on the virtual battlefield will be fully and continuously controlled by CFOR command entities.

The first generation of CFOR command entities has proven to be effective, but many key technical issues in modeling command decision making have not yet been addressed. For example, the suitability of the CFOR approach for modeling peer-to-peer communications or multi-level hierarchical C2 is untested. CFOR command entities currently interact with their peers in very limited circumstances using highly structured protocols to achieve specific goals (e.g., coordinate a passage of lines). They do not interact to negotiate for resources or otherwise collaborate during their decision processes. They also have a limited capability to convey or respond to planning guidance, the commander's intent, etc. Finally, current command entities include no mechanisms for varying the quality or nature of their decisions in ways consistent with the capabilities and limitations of human decision makers.

A primary goal of the Advanced Synthetic Forces effort in the DARPA Advanced Simulation Technology Thrust (ASTT) program is to develop technical solutions to these and other open issues in modeling human behavior in simulations. ASTT is seeking to develop technical innovations in modeling peer-to-peer communications, distributed/collaborative decision making, decision problems loosely constrained by doctrine, and realistic variability in behavior. The current set of funded efforts includes techniques for: modeling collaborative decision making by commanders and staffs; representing and reasoning about group structure (the roles and capabilities of group members) and group behavior (the goals and intentions of a group - friendly or opposing force - and the resulting implications for decision making); automated learning; and managing the loosely structured and flawed communications that can be expected between synthetic forces and human participants.

Human Behavior Modeling and Air Warrior Training

Dr. Dee H. Andrews

Decades of research in aircrew training has produced some fairly good information about how pilots behave in a variety of aviation settings. This research has been aimed primarily at normal and emergency procedural operations in the cockpit and so we have some reasonable data around which to build models and predictions about how pilots will perform these rote tasks. Even thought this data is available, the actual models that have been developed are relatively few. The models that have been developed have primarily been in the pilot selection domain. Considerably fewer models are available about how pilots learn these rote skills and therefore the aviation training community has been generally less influenced by these models than has the aviation selection community.

Cockpit Resource Management (CRM) is aimed at helping the aviation community understand how critical information is shared or not shared between aircrew members. CRM principles are taught to aircrews to help them communicate better. CRM is normally thought of in terms of multiplace aircraft crews, however the principles apply in large measure to fighter aircrews that work as a team even though they are in separate cockpits.

The last few years have produced some good empirical CRM data about how pilots communicate with each other in solving problems that afflict the flight deck. However, due partly to the relatively short time that CRM has been understood as a concept, the data about CRM is less abundant than is the case for data concerning how pilots behave towards operational and emergency procedures. This means the models available for CRM are fewer in number and are generally less robust than models for operational and emergency procedures. Since CRM has, at least anecdotally, been credited with reducing aircraft mishaps, and in some cases even saving life, a major effort in building quality models of pilot behavior in CRM situations is warranted.

From a warfighting training standpoint, very few valid models of pilot behavior in actual wartime settings have been developed. Such models are crucial if the military training community is to be able to take full advantage of the training capabilities that synthetic battlefields offer. Lack of quality models that accurately depict how aviation warfighters act when faced with tactical challenges means that the designers of large constructive wargame training exercises often make unrealistic assumptions about the behavior and capabilities of modeled aircrews, or the behavior of aircrews is ignored altogether. While this phenomenon is certainly not unique to the modeling of aviation warfighters behavior (this workshop has shown that the modeling of behavior for warfighters of all kinds is deficient), I contend that the effect of poor/absent modeling of aviation warfighters has particularly profound effects on constructive wargame training exercises due to the large impact that even a few aviation warfighters can have on a battlefield relative to their non-aviation counterparts.

The good news is that the cost of high fidelity, man-in-the-loop simulation is decreasing rapidly enough that it is possible to think of the day in the not too distant future when literally hundreds of high fidelity military aircraft simulators will be able to be linked into large training exercises. The synthetic battlefield will then have live pilots making real decisions as the battle scenario unfolds. Builders of synthetic training exercises will, in many cases, not have to worry about modeling pilots because the human aviators will be there. However, depending upon the size of the exercise, it will often still be necessary to model aviation warfighters because it is not likely that the military will ever have enough man-in-the-loop simulators to represent all of the necessary air warriors in very large exercises (e.g., theater level exercises). Also, there will be many instances when the analytical community will need to model pilot behavior without having to go to the trouble of bringing live pilots into their analyses.

Given the need to have quality, valid models of air warrior behaviors and decision making processes, how should the military proceed? First, we should do all that we can to conduct empirical research into understanding aircrew behavior. Synthetic battlefield man-in-the-loop simulators are becoming widespread enough, and their quality is now good enough, that many researchers in the DoD community should be able to have increasing access to these systems for measuring behavior in high fidelity battle scenarios. For example, at the Aircrew Training Research Division of the Air Force Research Laboratory we have just built a networked four ship simulation complex that includes four F-16C Block 30 simulator cockpits, with 360 degree high fidelity visual systems. This four ship simulation complex can allow a Flight Commander to have his four ship flight brief, fly and debrief together in a highly realistic setting. Mission tasks can be trained on this system that are not possible on training ranges due to safety, security, access and cost constraints. We plan to conduct a variety of studies with this four ship complex that will allow us to determine how a four ship flight behaves in combat settings. In addition, we are working for the Air Combat Command to develop a plan to examine the training effectiveness of networked F-15C four ship complexes that are being installed at Eglin AFB and Langley AFB in FY99. Part of the work we will perform with these complexes is to examine and measure teams in combat. Other efforts such as this are going on within the other services' aviation training research organizations.

Despite all the work that is either on-going or planned to collect empirical data that can be used to build and improve air warrior behavior models, there still is great need for analytical data collection. The main reason is the time it will take to collect enough empirical data for the modeling function. There are simply not enough researchers and research facilities available to collect the required data in a timely manner. Analytical efforts, primarily through interviews with aircrew subject matter experts, will be necessary to collect enough information for model construction in the relative short term. The need for air warrior behavior models is simply too pressing to wait for the empirical data collection efforts to pay dividends.

Finally, a great need exists for a systematic effort to build a quality data base of both empirical and analytical data that can be used for model building. Other DoD data bases have been established to catalog data for topics as diverse as ergonomics, geographical terrain and, logisitics operations. This same sort of effort and investment is needed to build a data catalog for human behavior modeling.

High Level Architecture for Modeling and Simulation

Dr. Duncan C. Miller

The Department of Defense (DoD), through the Defense Modeling and Simulation Office (DMSO) has developed and mandated the use of a common High Level Architecture (HLA) for Modeling and Simulation. The Undersecretary of Defense for Acquisition has designated the HLA as the technical architecture to be used in all simulations within DoD after 1999. The presentation will summarize the main tenets of the HLA and its implications for subsequent M&S development.

Another event of particular significance for the DoD M&S community is the emergence of a new, non-profit, independent standards organization, the Simulation Interoperability Standards Organization (SISO), which is currently in its final stages of formation. The presentation will outline the structure and key organizing concepts of SISO, including the semi-annual Simulation Interoperability Workshops (SIW), which have supplanted the previous Distributed Interactive Simulation (DIS) Workshops. The presentation will outline the plans for the Fall SIW on September 8-12 in Orlando.

Of particular relevance to the current audience is the SIW Forum on Human Decision-Making and Behavior Representation. The presentation will summarize the plans currently being finalized for this Forum's session at the September Workshop.

Rotorcraft Pilot's Associate

Mr. Keith Arthur

We enjoy a long and proud tradition of a well-equipped military. But in the cockpit of today's combat helicopters, being well-equipped means running headlong into two major problems: the cognitive and physical limits of the human. Simply put: the well-equipped cockpit provides too much information to digest in too short a time, and too many systems to control in too short a time. What is needed is a decision and action aid that also integrates the aircraft systems and mission equipment into a complete system.

The Rotorcraft Pilot's Associate program endeavors to do just that. The RPA program is an Advanced Technology Demonstration (ATD) whose objective is to provide the crew with an electronic "associate" and so to increase their combat mission effectiveness. Besides integrating an advanced mission equipment suite, RPA also monitors the mission progress, communications, and intelligence updates, considers the mission context, infers the crew's intent, and adapts its aid to the crew accordingly. RPA is composed of a data fusion module, a suite of planner modules, a suite of assessor modules, a task network, and a cockpit information manager.

Accounting for the human in this man-machine system is, of course, a critical part of the endeavor. RPA tackles this challenge with a three-pronged approach: a human-centered design philosophy, in which advanced and intuitive controls and displays are a key; a task network, part of the underlying software architecture; and a crew intent estimator, part of the cockpit information manager. We are discovering that, in practice, a human-centered design philosophy too easily slides back into a mission-centered philosophy, and sometimes even defaults back to a task-centered design philosophy. Of more interest here, though, are the latter two prongs of our approach. Our task network maps all of the tasks of an Army aviator onto a very large network of parallel and series tasks. So, no matter what the crew is doing at any particular point in any mission, there is a corresponding place on the task network. Contained in the task network are instructions that tell the associate's various modules how to then help the crew given the current mission context. The Crew Intent Estimator constantly monitors certain crew actions and certain aspects of the mission context to infer the crew's intent. If the inferred intent implies a goal different from the one RPA is working toward, the task network will "snap to" and synchronize the electronic associate with the human crew.

To give some meaning to this, consider the following operational scenario. On a reconnaissance mission, the helicopter suddenly gets painted by a threat radar. The advanced radar warning receiver recognizes the threat as a surface to air missile system in acquisition mode. The Battlefield Assessor module then triggers Actions of Contact, a thread of behavior designed to help the crew survive contact with the threat. The task network "snaps to" and as a result, an RF Hellfire missile is armed and actioned, the target acquisition system immediately locates and tracks the target, a masking location is found nearby to break line of sight to the threat, and visual and verbal cues are given to the pilot to help guide him to the masking location. The pilot, however, sees a better masking location and heads toward it. The Crew Intent Estimator recognizes that the pilot is going to a different masking location, and notifies the survivability planner, which immediately searches the digital map for a masking location in the direction the pilot is going. Visual and verbal cues are then modified to help guide the crew to their chosen masking location. En route the co-pilot squeezes off a missile which finds its target because RPA has activated the integrated flight and fire control system which brought the aircraft into weapon constraints even while maneuvering toward the masking location.

We are performing extensive developmental test and evaluation on RPA in manned simulation and plan to conduct a flight demonstration in late 1998 on a modified Longbow Apache.

In summary, our task network models the human in the scout/attack helicopter application, though coarsely. Experience so far has shown that this technique works very well in virtual simulation and we fully expect success in live simulation. The task network is a technique that may very well be transferable to the same and other applications in constructive simulation.

Simulations, Intelligence, and Simulated Intelligence

Mr. J. Darrell Morgeson

Early in 1993, LTG Wilson "Dutch" Shoffner gave the keynote address to a special meeting of the Military Operations Research Society (MORS) held at Ft. Leavenworth, Kansas. The theme of the conference focused on the development of C3I Measures of Effectiveness (MOE's) for warfighting simulations. Based on his experience as both an analyst and a field commander, General Shoffner made several points in his address that are still very relevant to the limitations of constructive warfighting simulations to adequately represent the human decision making or the C^nI process.

The essence of his argument was that certain key aspects of human reasoning were not capable of being emulated by computers using current state of the art approaches. These are called "command" functions, and are further characterized as the intuitive leaps that humans take in the inductive reasoning process. In the military vernacular, these are the processes that lead to the development of the commander's mental model of the battlefield. In contrast, "control" functions are those that are typically carried out by the staff, and are characterized by the deductive reasoning process. These functions do lend themselves to being represented by the computer and are often easily converted to algorithms, rules, and other decision support system functions. Taken together, the command and controls functions represent "intelligent reasoning" and the challenge for constructive simulations is how best to go about representing this phenomena.

Our work has been focused on developing alternative representational methods for assessing the effectiveness of alternative C^nI systems at the level of campaign warfare and below. Current approaches rely either on human-in-the-loop representations or expert systems approaches which seek to emulate the intelligent reasoning of commanders and their staffs. The former solution is expensive, limited in the number of alternatives that can be feasibly examined, and not repeatable. The latter is not sufficiently powerful to emulate the intelligent reasoning of humans. Accordingly, some of our new approaches seek to leverage the power of computer search and learning techniques to produce overall force behavior that is undifferentiable from the intelligent force behavior that commanders and staffs would have produced. Approaches derived from chess playing programs, artificial life methods such as genetic algorithms, and others are being explored.

We will present and example of how we have used these techniques to learn new tactics for the Airborne Laser Weapon's System. Time permitting, we will show a different example from the field of transportation. These techniques do not replace the traditional rule-based methods for representing C^nI process, rather they augment them. They are particularly applicable, when traditional methods break down -- when force asymmetries and completely new weapons concepts transcend our collective base of experience. They are methods of exploration and discovery, not tools of confirmation that "act out" the analyst's mental picture of the battlefield.

The research ongoing at Los Alamos seeks to provide a scientific and mathematical foundation for the appropriate use of these new techniques, in addition, to developing the algorithmic and computational techniques that will allow their effective use. The capability to use these tools for force structure and technology acquisition is far from being at hand, and will ultimately depend on teraflop computing resources of the next century.

Human Performance in Combat Simulation in Teams

Dr. A. J. Belyavin

Introduction

This presentation is based on an analysis of how the representation of human performance in combat simulation. can be improved. The development of a methodology for modelling human performance in systems was started in 1994 in the UK and this programme evolved into the development of the Integrated Performance Modelling Environment (IPME). The objectives of the development were threefold: to provide the basis for human performance elements for combat simulation models, provide a basis for modifying the behaviour of synthetic entities and provide the means for analysing the performance of individual systems while incorporating representation of human performance.

Human Performance in military OA

Human performance modelling in military operational analysis must meet the needs for representation, estimation and sensitivity. The representation must define the impact of the human element on the operational effectiveness of the system under study; the estimation of human performance or behaviour must include the effect of operational stressors; the level of detail in the human element should ideally include only those aspects which could affect the conclusions - sensitivity.

Performance modelling approach and issues

As a part of the specification for the IPME, the approaches to modelling the human operators in systems were reviewed and a number of potential predictive methodologies were examined. It was concluded that in modelling operator performance including teams, an environment using task analysis as the primary technique with a number of aspects of performance modelled either predictively or empirically, was the preferred approach. In a subsequent working paper, it was recommended that the environment should be based on MicroSAINT / HOS.

The approach adopted in this paper follows that used for IPME. It is assumed that in modelling human performance in a system, the semi-empirical task based approach should be adopted if feasible. Where this approach is impossible, due to task sharing between team members, it is possible to augment the approach with task scheduling algorithms. This line of attack satisfies the two basic requirements of representation and estimation. Sensitivity is an issue for the context in which performance is to be considered. There will be a need for each model or study to establish sensitivity to aspects of human performance.

The stochastic element in human performance has to be represented and, if simulation is the primary technique employed, rigorous experimental design has to be used for this stage of the study if costs are to be kept under control. A rigid application of the task network approach overrides the flexibility inherent in human planning and decision making. At the moment this flexibility can only be represented through goal directed models and substantial rule sets, and this has not yet formed part of the IPME project.

Task network simulation and data needs

The development of task network models for insertion in combat simulation needs data on the atomic elements within the task network, information on the effect of any stressors which are to be represented, and a sound description of the interactions between team members and the system. By using well defined generic atomic elements it is possible to restrict the need for specific data collection for a particular study. The effect of stressors can be extrapolated from a combination of laboratory studies and field data (where available) and can be applied to task times and errors in a systematic manner. The IPME makes assumptions about the taxonomic framework through which stressors affect task performance, defined in terms of execution time and probability of success. The task network logic flow expresses the consequences of both individual task failure and the time taken to complete a particular task. The output of the simulation is the time taken to complete the network and the probability of success.

Overall approach

The overall approach to the inclusion of human performance in a simulation can be summarised as follows:

Determine the components of the simulation which may be sensitive to the human elements

Define the task logic flow for those components including stressors, system elements and human elements

Simulate these "vignettes" using a model framework such as IPME, and produce results in a form appropriate to the simulation model in use

Incorporate the findings as data in the simulation model

Proof of principle

This approach has been tested in the UK in the last 12 months by generating launch response times for short range air defence systems for incorporation in an air attrition model. The simulation of response times for a system based on current UK practice will be briefly summarised and the results of the proof of principle study presented.

Unit Representation in Constructive Models

Dr. George R. Mastroianni

Constructive simulation represents the most challenging environment for representing human performance and human behavior. In virtual, live, or human-in-the-loop (HITL) simulations, humans provide some of the human performance needed for the purposes of the simulation. The designers of the system are relieved of the responsibility to represent, and to some degree, understand human performance because the human receives and processes inputs and produces outputs as a self-contained subcomponent of the system. In constructive simulations, on the other hand, human performance must be built into the system. The designers must know what aspects of human behavior must be represented, understand the functional relationships among simulation events, variables and outcomes, and develop a computational approach to implement this aspect of military operations into the simulation.

At Natick Research Development and Engineering Center, we have developed a constructive simulation called the *Integrated Unit Simulation System*, or IUSS. The IUSS is a comprehensive simulation environment emphasizing small-unit dismounted operations. Typical scenarios involve squad and platoon-sized elements. In the course of our work with this system, we have grappled with many issues that bedevil attempts to incorporate human performance in military simulations. During this presentation, I will describe the most important of these issues, discuss some of the approaches we have taken to solving these representation problems, and suggest key areas and approaches for future work.

The most difficult questions to answer about representing human performance in military simulation, in my opinion, are (1) What aspects of human performance need to be simulated?, and (2) With what fidelity do these aspects of human performance need to be simulated? How one answers these questions depends on the purposes for which the simulation is intended (especially, whether the simulation will have training or analysis as primary applications) the complexity of the simulation, the nature and diversity of the situations the simulation must be capable of representing, the nature and diversity of the simulated entities, and assumptions about the role of human performance in determining combat outcomes. In my experience, careful analysis of these questions is not the norm.

Once it has been decided what needs to be simulated, we must identify appropriate functional relationships among attributes of the simulated entities, environmental variables, task requirements, and tactical actions. For individual and small-unit dismounted operations, this is challenging. Sources of such information are research studies, records of performance in operational tests or realistic training exercises, doctrine, and the reports of experienced soldiers (subject matter experts). Information is often incomplete or contextually limited, but the time demands on system developers and the expense associated with acquiring new information often requires us to use what is available.

Finding a computational method to represent human performance in small-unit simulations can be more or less difficult depending on the complexity of the system and the particular aspect of performance to be represented. In general, current and past systems rely on performance degradation algorithms that are triggered by environmental conditions or task requirements. More recently, rule-based systems to represent tactical decision-making have been developed and implemented. In the future, artificial intelligence technologies such as fuzzy logic, adaptive learning, neural nets, and other innovative approaches may be recruited to represent more complex aspects of human behavior under more realistic conditions.

Overcoming the difficulties in representing human performance in constructive simulations will take time, effort, and ingenuity. The most important step, in my opinion, is to carefully analyze the requirements for human performance representation in each system. Because it is expensive and time-consuming to represent human performance, we must answer the "Why bother?" question clearly and convincingly early in the process.

Once we have decided what needs to be represented, we must approach the task of acquiring information energetically. We need to do a better job of mining existing sources of information and establishing methods to collect appropriate information from current and future operations and training. Resources tend to be directed more at system development than at the unglamorous, laborious, and difficult task of digging the foundation.

Finally, our experience has shown that there is more sensitivity to the problems of representing human performance when it is related to problems of decision making under uncertainty, situation awareness, suppression, and other cognitively complex situations. The more mundane aspects of human performance (locomotion, for example) are a more difficult sell as something deserving of time, attention, and money. We wouldn't think of building a model of combined arms combat without a good model of the vehicle dynamics of each of our tanks - we must be sure not to neglect the lower-level, "housekeeping functions" of human performance that may not be as intuitively appealing as decision-making but are nonetheless extremely important.

It will never be possible to insert a computational homunculus into a constructive simulation - that problem is intractable. What we can do, however, is sensibly circumscribe those aspects of human performance essential to our military simulation needs and then take advantage of what we already know to develop practical methods for representing those aspects of human performance in our simulations.

Affective Computing for Human Performance

Dr. Rosalind W. Picard

Most people realize that emotions are an important part of being human; emotions are powerful motivators, and can make the difference between the weary individual that quits and the one that presses onward to success. The theorist Silvan Tomkins argued that emotions are more powerful motivators than basic drives such as hunger, thirst and the need to breathe, citing how the lack of emotions, such as fear, which usually arise when a basic drive is unsatisfied, can lead to disastrous situations, as illustrated by pilots at 40,000 feet who did not wear oxygen masks or fear their inability to breathe, and subsequently met their deaths (Tomkins, 1962).

Whether or not Tomkins is right, there is compelling scientific evidence that emotions are not just powerful motivators, but essential to basic cognitive and RATIONAL functioning ---especially memory, rational decision making, perception, outlook on life, creativity, and more. Discussion and citation of studies supporting each of these can be found in Picard (1997). One set of surprising and very important findings is that of Antonio Damasio, who has studied patients who essentially have "too little" emotion because of a particular kind of brain damage. Instead of being highly rational, as one might expect since emotions are usually associated with irrationality, these patients are unable to behave in a rational manner. Damasio's findings illustrate that TOO LITTLE emotion impairs rationality, similar to what is already known for too much emotion. Healthy rational functioning requires a healthy balance of emotions (Damasio, 1994).

Basic emotional skills are critical for human intelligence. Psychologists have defined intelligence to include social and interpersonal intelligences -- especially skills such as the ability to recognize people's emotions, manage one's own emotions, experience and express empathy, and utilize emotions to motivate the achievement of goals (Gardner, 1983; Salovey and Mayer, 1990). Dan Goleman, in his book _Emotion Intelligence_ (1995) argues that such skills are even more important than IQ in determining an individual's success.

We cannot model intelligent human performance and interaction without modeling the mechanisms of emotion and their role in cognition, perception, decision making, motivation, and social interaction. A couple years ago I began a research program in "Affective Computing," computing that relates to, arises from, or deliberately influences emotions. The primary goal is the development of computers which exhibit emotional intelligence, for improving not only computer decision making, perception, and human-computer interaction, but also for furthering fundamental understanding about emotion and its influences.

Our initial efforts are focused on the development of computers which can recognize emotion, one of the building blocks of emotional intelligence. Toward this goal we have developed a prototype wearable affective computer, equipped with physiological sensors and special pattern recognition and signal processing algorithms. We are currently focusing on representation and recognition of emotions such as frustration, confusion, liking, disliking, and distress. One of the most difficult problems we face is the gathering of accurate data---expressions of genuine emotions, elicited under natural conditions, upon which to train recognition algorithms. Our expertise is in modeling, pattern recognition, signal processing, and algorithms for continuous learning, not in running human studies; we are teaming with experts in psychology to develop studies for gathering affective data. With the data collected to date, we have already achieved significant results in recognition of a small number of emotions in a person-dependent system. In some ways this work is like early research in speech recognition, where the initial results are dependent upon the speaker; however, in affect recognition it is significantly harder to gather data for training and testing new algorithms.

In my forthcoming book, I devote three chapters to illustrating various kinds of models of affect---low-level signal processing models, medium-level pattern models, and high-level rule-based models. These are intended to span the range of representations needed to capture both the physiological and cognitive-situational aspects of emotion. I describe dozens of specific theories and tools, both my own and those of others, which I believe will be important for modeling a broad scope of affective phenomena (Picard, 1997). The hope is that these tools will be immediately useful for researchers who are beginning to study human affect and to model it in computers.

References:

A. R. Damasio. Descartes' Error: Emotion, Reason, and the Human Brain. Gosset/Putnam Press, New York, NY, 1994.

H. Gardner. Frames of Mind. BasicBooks, New York, 1983.

D. Goleman. Emotional Intelligence. Bantam Books, 1995.

R. W. Picard. Affective Computing, MIT Press, To appear fall 1997.

P. Salovey and J. D. Mayer. Emotional intelligence. Imagination, Cognition and Personality, 9(3):185--211, 1990.

S. S. Tomkins. Affect, Imagery, Consciousness, volume 1: The Positive Affects. Springer, New York, 1962.

Entertainment Industry Research Directions and Inspirations

Dr. Michael Zyda

The Entertainment Industry has become large enough that its potential for funding research is close to that of many government agencies, particularly for computer graphics and the related technologies used for modeling and simulation. The National Research Council recently issued a report regarding this entitled "Modeling and Simulation: Linking Entertainment and Defense". In that report, a research agenda for joint defense and entertainment industry work was outlined. That agenda described the leading edge of modeling and simulation research for at least for the next twenty years. In this presentation, we discuss the research directions and inspirations possible from entertainment industry collaborations. Included in the discussion are:

Technologies for Immersion:

Image generationãgraphics computers capable of generating complex visual images.

Trackingãtechnologies for keeping track of the head position and orientation of participants in virtual environments.

Perambulationãtechnologies that allow participants to walk through virtual environments while experiencing hills, bumps, obstructions, etc.

Virtual presenceãtechnologies for providing a wide range sensory stimuli: visual, auditory, olfactory, vibrotactile and electrotactile

Networked Simulation:

Higher bandwidth networks--to allow faster communication of greater amounts of information among participants

Multicast and area of interest managers--to facilitate many-to-many communications while using limited bandwidth

Latency-reduction--techniques for reducing the true or perceived latency in distributed simulations

Standards for Interoperability:

Virtual reality transfer protocol--to facilitate large scale networking of distributed virtual environments

Architectures for interoperability--network and software architectures to allow scalability of distributed simulations without degrading performance

Interoperability standards--protocols that allow simulators to work together effectively and facilitate the construction of large simulations from existing subsystems.

Computer Generated Characters:

Adaptability--development of computer generated characters that can modify their behavior automatically over time

Individual behaviors-- computer-generated characters that accurately portray the actions and responses of individual participants in a simulation rather than those of aggregated entities.

Human representations--authentic avatars that look, move, and speak like humans.

Aggregation/deaggregation--the capability to aggregate smaller units into larger ones and deaggregate them back into smaller ones without sacrificing the fidelity of a simulation or frustrating attempts at interoperability

Spectator roles--ways of allowing observers into a simulation

Tools for Creating Simulated Environments:

Database generation and manipulation--tools for managing and storing information in large databases, to allow rapid retrieval of information, feature extraction, creation, and simplification.

Compositing--hardware and software packages that allow designers to form composite images with images taken from different sources (whether live-action footage or 3D models) and facilitate the addition or modification of lighting and environmental effects

Interactive tools--tools that use a variety of input devices (more than mouse and keyboard) to construct models and simulations.

Getting There from Here
Human Behavior in Existing and Future Simulations of Warfare

Dr. Thomas W. Mastaglio

Overview
The level of fidelity for all types of simulations has increased dramatically over the past decades. In the past, simulation users were content with combat outcomes that approximated the aggregate results of weapons employed by a force operating according to doctrine to accomplish their command and control training as well as analytic studies. However, improved computer interaction technologies that allow users better views of the battlefield have stimulated an appetite for increased fidelity and realism in these systems. This includes not only realistic scenes but also realistic performance of the iconic representations of the vehicles and soldiers which populate the battlefield. Representing individual vehicle has been tackled with rigor, and data accumulated over years of test and evaluation is being brought to bear in the computer software that accomplishes this. We now see that there exists a similar requirement for warriors, front line combatants as well as commanders and their staffs, to react and respond realistically. This in turn has led to recognition that we now need to increase the fidelity with which human behavior is represented.

Human Performance considerations in current simulations
A high level perspective of current simulations leads to a realization that human performance plays a greater role as one moves along a continuum from constructive to virtual to live simulations.

Constructive simulations today make little attempt to factor in the characteristics of the human. Their attrition equations could not be adapted to incorporate these factors, but this is not done because data is not available which tells us how combat outcomes are impacted by factors of human performance.

Virtual simulations have better fidelity of human performance than their constructive counterparts *because* humans are in the loop. The simulation algorithms do not have to concern themselves with representing humans, except those used for the semi-automated forces. Nevertheless, the overall outcomes of battles in virtual simulations still do not adequately capture the stress induced on these humans. Real warriors have to function on a battlefield where terrain, weather conditions, and the perception of real threat to life and limb can significantly mitigate performance. Users of virtual simulations are most often well-rested and comfortably located in either a sterile, air conditioned simulator, a mock-up operations center, or at a computer workstation.

Live simulations go the next step toward incorporating realistic human performance into the exercise because the humans are subject to the effects of the environment and other real world conditions. What is still missing though is the impact of real, not perceived, danger. Laser adjudicated warfare, except for the pressures associated with human competitive events, still cannot replicate the conditions that will stimulate battlefield conditions of danger.

What's Needed?
The challenge then is to devise a strategy and associated enabling information and technology which will support incorporating the effects of human performance variability into the predicted outcomes of the constructive, virtual, or live simulations. The outcome from this workshop should be structured to establish requirements or identify existing capabilities as seen in three areas.

Theoretic structures are required. These are most likely to be useful science from psychology. We have to have an underlying scientific basis for the human behavioral models that will be used in future simulations. A National Research Council Study is focusing on this issue.

Empirical results are needed to drive the simulation. The simulation technical community needs data from actual combat, or other experiments which can be used within a theoretic framework to predict the results of behavioral factors on overall individual, unit, and operational performance. An ongoing Army Science Board Study is trying to identify existing data sources and deficiencies in understanding human behavior for Army needs – to include doctrine, training, leadership and simulation development.

Algorithmic approaches that simulationists can implement with computer software need to be identified. I recommend that this be the primary focus of the working group discussions and that we either leave this workshop with agreement as to what will work or a recommendation for research & development needed. What is possible today and in the near future should in turn determine requirements for the theoretic structures and for the data to drive those algorithms. The simulation community needs to articulate its needs, but do so within the context of what they feel can be algorithmically applied to increase the fidelity of human performance representations. It is an appropriate time in the evolution of simulation technology to tackle this problem as a community while engaging the theoreticians and data gatherers to meet our needs.